Niveau(x) d'enseignement: 4 Numéro EDUQ : 8570872
Centre d'intérêt: 874

Cloutier, Jean-François.

Elaboration et expérimentation d'un modèle algorithmique comme support à l'apprentissage de concepts technologiques abstraits et leur transfert en situations concrètes chez des étudiants du secteur professionnel court / J.-F. Cloutier, B. F. Rousse, H. D. Stolovitch. -- Montréal : Université de Montréal, Programme de perfectionnement des maîtres en enseignement professionnel, 1985.
vi, 75 f. ; 28 cm.

Ss titre: Rapport sur le développement d'un prototype de progiciel d'aide à l'analyse cohérente des réseaux conceptuels "ARC-0.1". "Rapport final III (82/MPS/02-PPMEP 03)". Résumé: f. 01-02.

Descripteur(s): *Techniques physiques *Concept *Enseignement assisté par ordinateur *Evaluation du matériel *Informatique *Secondaire (Niveau)

Identificateur(s) de sujet: *Apprentissage de concepts *Analyse des concepts *Intelligence artificielle *Progiciel

---

Le rapport rend compte des travaux de développement d'un pré-prototype de démonstration (ARC 0.1) visant à la création d'une méthodologie générale d'analyse des réseaux conceptuels et à l'étude de faisabilité d'un modèle informatisé permettant d'aider les "experts" de n'importe quel domaine de connaissance à réaliser des analyses conceptuelles sous forme de réseaux de façon cohérente et rigoureuse.

Le rapport identifie, dans un premier temps, la faiblesse du type de représentation des concepts comme cause principale des limitations du prototype algorithmique qui avait pour mission de valider et d'expérimenter à grande échelle une méthode didactique conceptuelle. Il propose ensuite un formalisme de représentation des concepts en réseaux cohérents qui permettra de surmonter ces limitations. Enfin, la nécessité d'un programme interactif et intelligent d'aide à l'analyse de concepts est discutée et apporte un argument supplémentaire en faveur du développement du formalisme de représentation.

Le développement d'un formalisme de représentation des concepts est l'étape première et fondamentale de tout système informatique espérant traiter de l'information conceptuelle de façon adaptive et intelligente. Cette étape est en voie d'être franchie. La mise au point d'un système d'aide à l'analyse des concepts est maintenant une entreprise théoriquement possible et déjà bien amorcée. Pour la mener à bien, il faudra néanmoins compter sur un effort soutenu d'une durée que l'on peut estimer à deux ans.

Type de document: r
Financement: Bénéficiaire - Université de Montréal
Organisme(s) source(s) - MEQ-DGERU
Programme(s) de subvention - PPMEP

EDUQ 8570872

ELABORATION & EXPERIMENTATION D'UN MODELE ALGORITHMIQUE COMME SUPPORT A L'APPRENTISSAGE DE CONCEPTS TECHNOLOGIQUES ABSTRAITS ET LEUR TRANSFERT EN SITUATIONS CONCRETES CHEZ DES ETUDIANTS DU SECTEUR PROFESSIONNEL COURT

(RAPPORT SUR LE DEVELOPPEMENT D'UN PROTOTYPE DE PROGICIEL D'AIDE A L'ANALYSE COHERENTE DES RESEAUX CONCEPTUELS "ARC-0.1")
(Janvier 85-Mai 85)

CLOUTIER, J-F.
ROUSSE, B.F.
STOLOVITCH, H.D.

AOUT 1985

Rapport Final III

(82/MPS/02-PPMEP 03)

PROGRAMME DE PERFECTIONNEMENT DES MAITRES
EN ENSEIGNEMENT PROFESSIONNEL

UNIVERSITE DE MONTREAL * FACULTE DES SCIENCES DE L'EDUCATION

ARC.TIT: 11/08/85

EDUQ 8570872

ELABORATION & EXPERIMENTATION D'UN MODELE ALGORITHMIQUE COMME SUPPORT A L'APPRENTISSAGE DE CONCEPTS TECHNOLOGIQUES ABSTRAITS ET LEUR TRANSFERT EN SITUATIONS CONCRETES CHEZ DES ETUDIANTS DU SECTEUR PROFESSIONNEL COURT

(RAPPORT SUR LE DEVELOPPEMENT D'UN PROTOTYPE DE PROGICIEL D'AIDE A L'ANALYSE COHERENTE DES RESEAUX CONCEPTUELS "ARC-0.1")
(Janvier 85-Mai 85)

CLOUTIER, J-F.
ROUSSE, D.F.
STOLOVITCH, H.D.

AOUT 1985

Rapport Final III

(82/MPS/02-PPMEP 03)

PROGRAMME DE PERFECTIONNEMENT DES MAITRES
EN ENSEIGNEMENT PROFESSIONNEL

UNIVERSITE DE MONTREAL * FACULTE DES SCIENCES DE L'EDUCATION

ARC.TIT: 11/08/85

NOTE

Ce texte fait partie des documents produits dans le cadre du projet de recherche PPMEP(82/MPS/02), réalisé entre janvier 1983 et mai 1985, ayant pour titre:

"ELABORATION & EXPERIMENTATION D'UN MODELE DIDACTIQUE INFORMATISE POUR L'APPRENTISSAGE DE CONCEPTS TECHNOLOGIQUES ABSTRAITS CHEZ DES ETUDIANTS DU SECTEUR PROFESSIONNEL".

Cette recherche a pu être menée grâce à une subvention du Ministère de l'Education du Québec (PPMEP).

Ce projet a été supervisé et dirigé par:

. Harold STOLOVITCH, *Directeur du projet et Professeur agrégé en Technologie éducationnelle, Université de Montréal.*

. Viateur LEMIRE, *co-directeur du projet et Responsable du programme PPMEP, Faculté des Sciences de l'éducation, Université de Montréal.*

A:NOTEDIT

Table des matières

Sommaire i

Introduction iii

I- Résumé. 01

II- Base de données & Algorithmes v/s Base de Connaissances & Raisonnement automatique . 03

III- Le Concept de Formalisme. 10

IV- Réseaux cohérents & Caractéristiques minimales d'un Formalisme de Représentation 13

- Information contenue dans un concept et une relation. 13
- Réseaux conceptuels et relationnels 15
- Caractéristiques mininales d'un formalisme de Représentation. 17

V- Eléments de Base du Formalisme de Représentation. . 21

- Vers le Formalisme. 26
- Concepts. 29
- Relations 37

VI- Description complète du Formalisme de Représentation 42

- Paires ordonnées, Intervalles temporels de validité & Graphes. 42
- Projections & Descripteurs. 44

- Relations spéciales 50

- Connecteurs logiques, Régions d'un Graphe & Lignes d'Identité 52

- Paires ordonnées & Graphes comme Eléments de Paires ordonnées. 57

- Opérations de Spécialisation & de Généralisation. 61

VII- Dynamisme de la Création & de la Modification des Réseaux 69

VIII- ARC 0.1, un Pré-Prototype & Prospectives 73

* * * * *

INTRODUCTION

Le présent rapport rend compte des travaux de développement d'un pré-Prototype de démonstration (*ARC 0.1*), visant à la création d'une méthodologie générale d'Analyse des Réseaux Conceptuels et à l'étude de faisabilité d'un modèle informatisé permettant d'aider les "*experts*" de N'IMPORTE QUEL DOMAINE DE CONNAISSANCE à réaliser des analyses conceptuelles, sous forme de réseaux et de façon "*cohérente*" et "*rigoureuse*".

Les résultats obtenus jusqu'à présent tendent à démontrer qu'un formalisme de représentation des concepts en réseaux, à la fois cohérent et rigoureux, est possible en s'appuyant sur les données théoriques de la logique propositionnelle. De plus, l'universalité d'application du modèle à TOUT DOMAINE DE CONNAISSANCE, qui représente une condition "*sine qua non*" pour l'opérationnalistion effective d'un Tuteur Conceptuel Universel du genre de SOCRATIX (*cf. Rapports I, II et al.*), est rendue formellement possible du fait qu'il se fonde sur les structures formelles des "*règles de clasification*" et des "*réseaux conceptuels*" plutôt que sur les contenus spécifiques des concepts.

En fait, les travaux théoriques menés jusqu'à maintenant, dans le cours de ce projet de recherche, permettent de démontrer, non seulement que la réalisation d'un "*système-expert*" permettant de faciliter l'analyse cohérente et rigoureuse de "*réseaux conceptuels*" est possible, avec les ressources actuelles en Intelligence Artificielle, mais encore qu'un tel système serait applicable à TOUT DOMAINE DE CONNAISSANCE STRUCTUREE, et que, par ailleurs, l'échec relatif des recherches menées antérieurement et se fondant sur les "*réseaux sémantiques*" est lié précisément à l'incohérence et à au manque de rigueur formelle de ces mêmes réseaux.

Toutefois, la nouvelle voie ouverte dans le domaine de l'analyse des concepts et des réseaux conceptuels, si riche et prometteuse qu'elle soit, ne saurait encore être considérée comme un modèle théorique pleinement finalisé et effectif: dans l'état actuel, il ne s'agit encore que d'une tentative, sérieuse certes, mais dont les fondements ne constituent qu'une base pour des recherches de développement ultérieurs considérables.

La finalisation d'un modèle universel d'analyse de réseaux conceptuels, dont les avantages seraient énormes et innombrables au-delà même de la stricte formation professionnelle et technologique (*rationnalisation des banques d'information, optimalisation de la recherche scientifique, amélioration de la compétence des ingénieurs et experts de ious domaines, accroissement de la rigueur sémantique dans la pratique linguistique, etc..*), représente un intérêt évident.

Or, compte tenu d'une concentration adéquate de chercheurs de haut-niveau, d'un équipement de développement informatique suffisant et d'une concertation inter-institutionnelle effective, nous pensons qu'un tel projet majeur - incluant la mise au point totale d'un Tuteur Conceptuel Universel (*SOCRATIX*) et la constitution d'une banque variée de contenus cohérents validés, dans de multiples domaines de formation - est parfaitement réalisable dans un délai de cinq années, soit d'ici 1990.

Conscients du fait qu'un tel ensemble d'outils est fondamental, face aux problèmes cognitifs rencontrés dans la formation professionnele et technologique, et que, de plus, l'impact en serait international et situerait le Québec dans une position de pointe en regard respectivement de la recherche didactique et de l'intelligence artificielle, nous osons espérer qu'un tel projet pourra être effectivement concrétisé, plutôt que de demeurer un "*voeux pieux*" en attendant que d'autres le réalisent, AILLEURS!

B.F.ROUSSE, Ph.D.

* * * * *

RESEAUX CONCEPTUELS COHERENTS:

DEVELOPPEMENT D'UN FORMALISME DE REPRESENTATION

I. RESUME

Une première étape du projet d'informatisation de Socratix, un tuteur conceptuel intelligent et universel, fut la construction d'un prototype algorithmique qui avait pour mission de valider et d'expérimenter à grande échelle une méthode didactique conceptuelle. De plus, le prototype a permis de démontrer concrètement les limitations inhérentes à l'approche algorithmique. (*Le Rapport Technique no.IV/ P.P.M.E.P.-10 fait état de ces limitations*).

Le présent rapport identifie la faiblesse du type de représentation des concepts comme cause principale des limitations du prototype. Il propose ensuite un formalisme de représentation des concepts en réseaux cohérents qui permettra de surmonter ces limitations.

Un programme (*ARC*) a été mis au point, afin de démontrer la potentialité du formalisme de représentation. Bien qu'il n'incorpore pas toute la complexité du formalisme, ce démonstrateur est néanmoins capable d'organiser en réseau cohérent hiérarchique un ensemble de concepts prédéfinis. Il accomplit sa tache en RAISONNANT à partir des définitions des différents concepts.

La nécessité d'un programme interactif et intelligent d'aide à l'analyse de concepts est discutée et apporte un argument supplémentaire en faveur du développement du formalisme de représentation. Plusieurs propriétés du formalisme se prêtent fort bien au problème de guider intelligemment l'analyse de concepts qui est essentiellement un processus créatif rigoureux.

* * * * *

II. BASE DE DONNEES ET ALGORITHMES VERSUS BASE DE CONNAISSANCES ET RAISONNEMENT AUTOMATIQUE.

Une base de données et une base de connaissances sont toutes deux des structures organisant des informations codifiées. Elles diffèrent par la façon dont elles structurent ces informations et conséquemment par les opérations de traitement de l'information qu'elles supportent.

Essentiellement, une base de données organise ses informations afin de permettre un accès efficace (*en terme de temps d'exécution*) à ses différentes constituantes. L'opération principale est donc d'accèder à chaque information et la seule propriété que DOIT posséder la structure organisant les informations est cette possibilité d'accèder à chaque information.

Les différents algorithmes qui déterminent les mécanismes d'accès ne font qu'UTILISER les propriétés structurelles de la base de données. Par exemple, une structure en arbre ordonné permet un algorithme d'accès plus efficace à chaque information qu'une structure simplement séquentielle et ordonnée.

Le prototype Socratix 0.2 possède une base de données qui organise les informations décrivant les différents concepts. Ceci est tout à fait adéquat pour permettre l'accès à la définition d'un concept, à la description graphique des composants de ses séries rationelles ou à d'autres informations directement codées. Les problèmes apparaissent lorsqu'on cherche un autre type d'information.

Prenons la situation suivante: Pour faire le diagnostique d'un mauvais apprentissage d'un concept, il est nécessaire de vérifier l'apprentissage correct des concepts qui sont utilisés pour énoncer les attributs de ce concept.

Par exemple, quand Socratix 0.2 décèle que son interlocuteur n'a pas appris correctement que dans un circuit parallèle "il n'y a qu'une résistance par trajet", c'est tout ce que le prototype peut affirmer. Il ne peut émettre l'hypothèse que le problème réside peut-être dans un mauvais apprentissage du concept de "trajet". Dans la base de données, l'attribut "il n'y a qu'une résistance par trajet" est la deuxième chaine de caractères dans l'ensemble des attributs de "circuit parallèle", et rien de plus. Le prototype n'a pas accès au fait que des concepts composent l'énoncé.

Ce qui manque à Socratix 0.2 c'est de l'information sur la façon dont les concepts sont organisés ENTRE EUX. La base de données ne contient aucunement ce type d'information.

Il faut donc une organisation de l'information qui contienne les PROPRIETES STRUCTURELLES des concepts. Quelles sont ces propriétés structurelles?

Un concept est déjà une structure en soi. Il est une série d'attributs qui décrivent les conditions minimales d'appartenance à la classe définie par le concept. C'est donc une règle de classification. Chaque attribut contient dans son énoncé d'autres concepts: ceci constitue la dimension HORIZONTALE d'un réseau de concepts. Chaque concept possède au moins une SURCLASSE qui est un concept plus général (*ou un ensemble de SOUS-CLASSES qui sont des concepts plus spécifiques*): ceci constitue la dimension VERTICALE ou HIERARCHIQUE du réseau conceptuel.

Ainsi chaque concept est défini par d'autres concepts et est un cas spécialisé d'au moins un autre concept. La signification d'un concept dépend donc de la signification de beaucoup d'autres concepts. De façon figurative, on peut affirmer que la signification d'un concept est déterminée par sa POSITION RELATIVE dans un réseau de concepts.

Le concept est un type de structuration de la connaissance et l'investigation des propriétés structurelles des concepts relève de la recherche épistémologique. Trouver un moyen de représenter les concepts dans un ordinateur afin de recréer ces mêmes propriétés structurelles, sur lesquelles des programmes raisonneront, relève de la recherche en Intelligence Artificielle (*I.A.*). L'I.A. peut être succintement décrite comme une forme d'épistémologie appliquée.

Une base de connaissances est une organisation d'informations qui se veut la plus fidèle possible aux structures organisationelles de la connaissance telle que nous la "connaissons". Le critère de structuration n'en est plus un d'efficacité d'accès (*base de données*) mais de fidélité aux propriétés informationelles de la connaissance (***base de connaissances***).

Le passage d'une base de données à un base de connaissances est nécessaire à la réalisation de l'objectif d'universalité du tuteur. Le tuteur doit pouvoir enseigner n'importe quel domaine de contenu conceptuel. Il ne doit donc pas être un spécialiste dans un domaine précis. Peut-il enseigner un contenu sans être un spécialiste de ce contenu?

Il faut d'abord décomposer l'habilité d'enseigner en un ensemble d'habilités et inspecter les stratégies qui décrivent chacune. Pour enseigner un contenu, il faut avoir une méthode didactique qui voit aux stratégies de transfert de l'information. Il faut une méthode diagnostique qui vérifie la qualité du transfert et finalement une expertise pédagogique qui décide des meilleures modalités possibles pour effectuer le transfert.

Examinons des stratégies qui décrivent l'expertise pédagogique:

1- *Avant d'aborder l'enseignement d'un concept, il faut s'assurer que les concepts pré-requis aient été "suffisamment" appris.*

2- *On n'enseigne pas plus que trois niveaux conceptuels lors d'une même leçon (verticaux ou horizontaux).*

3- *Si un manque de motivation apparaît après une excellente performance, il faut diminuer la redondance dans la présentation du contenu.*

Aucune de ces stratégies n'est spécifique à un domaine de contenu quelconque. En fait les stratégies 1 et 2 s'expriment en fonction de propriétés structurelles du réseau de concepts pour le domaine de contenu enseigné. Les prérequis d'un concept sont tous les concepts qui, directement ou indirectement, participent à la définition du concept.

On a enseigné trois niveaux conceptuels horizontaux lorsqu'on a enseigné un concept, un de ses pré-requis et un des pré-requis de ses pré-requis.

L'application de ces stratégies demande au tuteur de raisonner sur les propriétés structurelles d'un réseau, d'où la nécessité d'une représentation des concepts adéquate. Ces règles s'appliquent à l'enseignement de tout contenu exprimé en réseaux conceptuels, et tous les réseaux conceptuels ont les propriétés structurelles qui sont le substrat des stratégies d'enseignement des concepts.

Un tuteur dont les stratégies se basent sur des propriétés générales à tout contenu conceptuel sera donc véritablement UNIVERSEL.

Pour que le tuteur soit intelligent, il faut que les règles qui dictent son comportement lui permettent une bonne mesure d'adaptivité et garantissent un minimum d'efficacité. Ceci n'est pas le sujet du présent rapport. Par contre, si l'information de contenu disponible au tuteur ne permet pas d'énoncer des règles à caractère universel, il faut abandonner le projet d'un tuteur universel ou même intelligent.

Ayant démontré la nécessité d'un mode de représentation appellé base de connaissance, les prochaines sections seront consacrées à la description du formalisme de représentation des concepts.

* * * * *

III. LE CONCEPT DE FORMALISME.

Un formalisme est un attachement à la forme et non au contenu. Prenons comme exemple de formalisme la logique propositionelle.

La logique propositionelle examine uniquement la forme d'un discours constitué de propositions et n'est pas du tout concernée par son contenu. Une proposition n'a comme propriété que d'être vraie ou fausse.

Ces trois propositions ont une forme identique:

1. *Il pleut en Zambie ET il grêle en Alaska*

2. *J'aime le cognac ET la vie est douce*

3. *A ET B*

La logique propositionelle décèle la contradiction exprimée par chacun de ces deux ensembles de propositions:

1. *Tous les hommes sont mortels;*
   Socrate est un homme;
   Socrate n'est pas mortel

2. $\forall x\ A(x) \Rightarrow B(x)$
   $A(c)$
   $\neg B(c)$

Il n'est pas nécessaire de connaître Socrate pour affirmer que le premier ensemble de propositions est contradictoire; un examen logique de la FORME des deux discours suffit à établir de façon identique les contradictions.

La logique propositionelle n'est donc pas concernée par le sens des symboles, ou sémantique (*à l'exception des connecteurs logiques tels que "et", "ou", "implique" etc...*).

Utilisons maintenant le langage PROLOG comme exemple d'un formalisme général qui permet le traitement automatique de contenus spécifiques.

Le langage de programmation PROLOG utilise les règles de la logique propositionelle et un algorithme automatique de preuve par contradiction (*résolution*). Il permet d'investiguer les conséquences logiques de propositions.

Son utilisation la plus courante est dans la construction de systèmes experts qui sont des spécialistes informatisés. On peut décrire en PROLOG un système expert en mécanique automobile sans que le langage PROLOG n'ait préalablement de connaissances en mécanique.

Exemple d'une section de programme en PROLOG:

.
.
.

brule(x) if moteur(x) & chauffe(x) & manque_d'huile(x)
chauffe(x) if moteur(x) & RPM(x y) & plus_grand(y 8000)

.
.
.

PROLOG est un langage formel de représentation qui est indépendant de tout contenu et qui peut traiter une expértise dans n'importe quel domaine s'exprimant en propositions logiques.

Toute cette discussion vise à démontrer qu'il est possible à un contenu spécifique d'être représenté dans un formalisme qui est indépendant de tout contenu, formalisme qui permet pourtant de manipuler ce contenu.

Parallèlement, il est possible (*et nécessaire*) de représenter tout domaine de contenu conceptuel par un formalisme indépendant de tout contenu spécifique et permettant d'extraire des concepts ainsi représentés toute l'information contenue dans un réseau conceptuel (*attributs, prérequis, surclasses, "distance conceptuelle" entre deux concepts, etc.*).

* * * * *

IV. RESEAUX COHERENTS & CARACTERISTIQUES MINIMALES D'UN FORMALISME DE REPRESENTATION.

A. Information contenue dans un concept et une relation

Voyons d'abord quelle information possède un concept lorsqu'il est en réseau.

Prenons le concept de TRIANGLE ISOCELE:

"Un triangle isocèle est un triangle dont deux des côtés sont de même longueur."

Après une analyse de l'énoncé on dégage l'information suivante:

SURCLASSE = *TRIANGLE*

ATTRIBUT : - *deux côtés ont une longueur identique*

On remarque que l'attribut contient le concept de côté (*x est un côtéPN*) *et la relation longueur* (*x est la longueur de y*). Il n'est pas nécessaire d'inclure le fait qu'un triangle isocèle a trois côtés. Cette information est accessible dans le concept de triangle, sa surclasse.

Voyons maintenant qu'elle information est contenue dans la définition d'une relation.

Prenons la relation CONNECTE-A:

"Un être est connecté à un autre être si une des composantes du premier être est contigu à une des composantes de l'autre être. La relation connecté-à est un type de lien."

Après une analyse de l'énoncé on dégage l'information suivante:

SURCLASSE : *lien*

ATTRIBUTS : *-le premier être a des composantes*
-le deuxième être a des composantes
-certaines composantes du premier être sont contigues à certaines composantes du deuxième être

On remarque donc que la relation CONNECTE-A utilise le concept être (*x est un être*) et les relations composante (*x est une composante de y*) et contigue (*x est contigue à y*).

Ainsi jusqu'à présent nous n'avions mentionné en fait de réseaux que les réseaux conceptuels. Ceci était incomplet. La discussion précédente démontre qu'un concept contient des attributs et que ces attributs sont exprimés en termes de RELATIONS.

Or tout comme un concept, une relation entre deux êtres se définit par un ensemble d'attributs. De plus une relation peut avoir plusieurs surclasses relationelles dont elle est un cas spécialisé. les relations s'organisent donc elles aussi en réseau et pour qu'un réseau relationel soit valide, il faut aussi que les définitions qui la composent soient cohérentes entre elles.

Nous abordons de façon plus formelle la description des réseaux conceptuels et relationels ainsi que leurs représentations.

B. Réseaux conceptuels et relationels

Un réseau conceptuel (*ou relationel*) est un ensemble cohérent de concepts (*ou relations*). Les définitions doivent être cohérentes entre elles et chaque attribut doit en plus faire un emploi cohérent des relations et concepts le composant. (*Les règles décrivant les contraintes de cohérence seront abordées plus loin*).

Une structure conceptuelle (*ou conceptPN*) est constituée des noms des concepts qui sont les surclasses directes du concept, et d'un ensemble d'attributs valides définissant les conditions minimales d'appartenance à la classe d'êtres décrite par le concept.

Pour valider l'inclusion d'un être dans une classe décrite par un concept, il faut valider d'abord son inclusion dans toutes les surclasses directes du concept. Ensuite on doit s'assurer que l'être respecte tous les attributs du concept. Le processus est <u>récursif</u> et nous verrons plus tard les conditions assurant sa terminaison (par le principe d'un horizon conceptuel).

Un attribut est une relation ou une composition de relations qui est établie entre des êtres par observation à un moment quelconque. Une relation est définie (*dans un réseau relationel*) en termes d'attributs concernant les deux êtres impliqués dans cette relation.

Comme pour les concepts, les relations sont organisées en réseau bi-dimensionel. La dimension horizontale vient du fait qu'une relation est souvent décrite par un ensemble de relations conjointement vérifiées (*attributs*). La dimension verticale vient du fait qu'une relation peut avoir des surclasses; une relation peut être un cas spécialisé d'une relation plus générale.

Par exemple, la relation FRERE-DE est un cas spécial de la relation PARENT-DE impliquant deux êtres éléments de la classe décrite par le concept ETRE-VIVANT.

Un concept ne doit pas contenir d'information redondante; les attributs ne doivent pas répéter inutilement des attributs d'une surclasse (*directe ou indirecte*). En fait, les attributs doivent décrire en quoi le concept est une forme SPECIALISEE de ses surclasses directes.

Par exemple, un quadrilatère a pour surclasse POLYGONE et comme attribut le fait que son nombre de côtés est égal à 4.

La même discussion s'applique aux réseaux relationels.

C. Caractéristiques minimales d'un formalisme de représentation

Un formalisme de représentation des concepts en réseaux doit pouvoir décrire tout concept ou relation de façon à indiquer son appartenance à un réseau, doit permettre l'application de méthodes automatiques de vérification de la cohérence du concept relativement à l'information déjà contenue dans le réseau et surtout doit permettre de contenir de façon non-ambigue la sémantique du concept. Examinons chacune de ces caractéristiques.

La première caractéristique du formalisme est de permettre d'indiquer les surclasses d'un concept (*il suffit simplement de les nommer*) et de rendre explicite l'emploi de concepts et de relations dans la définition de chaque attribut.

Il faut ensuite que le formalisme permette la vérification automatique de la cohérence d'un concept relativement au réseau. Voyons donc quelles sont les contraintes de cohérence qu'il faut vérifier.

Chaque attribut doit être valide en soi. Vérifier la validité d'un attribut revient à s'assurer que l'utilisation des relations et concepts dans son énoncé se fait de façon cohérente avec leurs définitions contenues dans les réseaux. (*D'autres contraintes de validité d'un attribut seront abordées lors de la description des éléments du formalisme de représentation*).

Après avoir vérifié la validité individuelle de chaque attribut, il faut s'assurer que l'ENSEMBLE des attributs est cohérent avec les définitions des surclasses du concept. Puisque le concept est un cas spécialisé de ses surclasses, il est logique de considérer certains de ses attributs comme des formes spécialisées d'attributs des surclasses.

Prenons un attribut de triangle isocèle qui dit "au moins deux côtés sont de longueur égale". Dans le concept de triangle équilatéral (*une sousclasse de triangle isocèle*), cet attribut est spécialisé pour devenir "trois côtés sont de longueur égale".

On peut donc percevoir un attribut d'un concept comme le résultat d'une OPERATION de spécialisation effectuée sur les définitions de ses surclasses (*spécialisation verticale*) ou sur les définitions de concepts et/ou relations participant à sa définition (*spécialisation horizontale*).

Nous définirons plus loin l'ensemble des opérations qui permettent de générer des attributs par spécialisation.

Les autres attributs du concept sont de l'information nouvelle ne provenant d'aucune définition de concept ou relation (*spécialisation libre*).

Conséquemment, nous pouvons proposer qu'un attribut est valide s'il est le résultat correct d'une opération de spécialisation définie.

Notre formalisme doit donc permettre la définition d'opérations de spécialisation comme générateurs d'attributs afin de rendre possible une vérification de la cohérence des réseaux (*qui est la deuxième caractéristique essentielle du formalisme*).

La troisième caractéristique essentielle du formalisme est d'être capable de contenir de façon non-ambigue la sémantique d'un concept. Il est aussi fortement souhaitable que ce formalisme puisse exprimer la sémantique du plus grand nombre possible de concepts . Il faut ensuite que des opérations formelles sur des attributs génèrent de nouveaux attributs qui eux aussi aient un contenu sémantique correct.

Pour démontrer que le formalisme possède cette caractéristique, il nous faudra présenter les éléments du formalisme pour ensuite illustrer de façon empirique la généralité de son pouvoir représentationel.

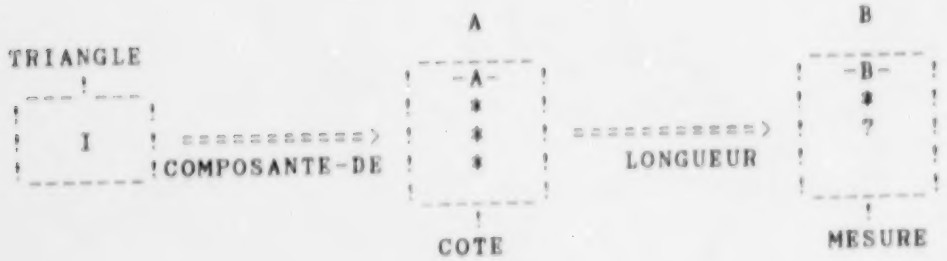
* * * * *

V. ELEMENTS DE BASE DU FORMALISME DE REPRESENTATION.

Le formalisme utilise les concepts d'ensemble, de projections d'un ensemble sur un autre (*"mapping"*), de partition et de construction d'ensembles.

Avant de présenter les définitions, voici un exemple qui donnera une appréciation intuitive de quelques éléments du formalisme.

L'attribut du concept triangle isocèle "il y a au moins deux cotés de même longueur" se représente comme suit:



ou

I (= 1) : composante --> A (= 3)
A (= 3) : longueur --> B (<1..2>)

L'ensemble I contient l'être que l'on cherche à identifier comme triangle isocèle. L'ensemble A contient les composantes de cet être qui doivent être au nombre de trois. L'ensemble B contient les valeurs des longueurs des différentes composantes; il y a AU PLUS DEUX valeurs différentes.

Dans cet exemple, tous les ensembles sont DELIMITES; l'ensemble "A" ne doit contenir que des éléments qui sont des COTES, tandis que l'ensemble "B" ne peut contenir que des éléments qui sont des MESURES.

Quant à "I" son élément doit être un TRIANGLE; il doit donc appartenir à la surclasse.

Les flèches dans les diagrammes représentent les projections d'un DOMAINE à un CO-DOMAINE. Une projection est en soi un ENSEMBLE DE RELATIONS identifiées par le nom de la projection. Par définition, une projection d'un ensemble sur un autre implique que tous les éléments du DOMAINE sont reliés à au moins un élément du CO-DOMAINE et que tous les éléments du CO-DOMAINE sont reliés à au moins un élément du DOMAINE.

Plusieurs informations sont associées à chaque projection. Son nom indique l'identité de l'ensemble des relations qui composent la projection. Ces informations sont les seules présentes dans l'exemple ci-haut.

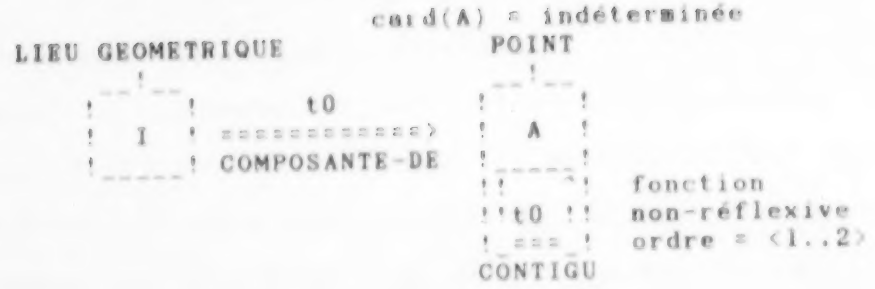
On peut ensuite indiquer si la projection est de type fonction (*deux êtres du domaine ne peuvent être reliés exactement au même sous-ensemble d'éléments du co-domaine*). On peut donner l'ordre de la projection (*combien de relations unissent chaque élément du domaine à des éléments du co-domaine*). De plus, on peut préciser A QUEL MOMENT la projection doit être vérifiée relativement à l'intervalle d'identification, intervalle durant lequel l'appartenance de l'être à la classe définie par le concept est validée (*ceci sera abordé plus loin*).

A chaque ensemble, on peut associer de l'information concernant sa cardinalité (*le nombre de ses éléments*).

Toutes ces informations peuvent être données avec différents niveaux de précision. Augmenter cette précision est un moyen de générer des attributs plus spécifiques.

Prenons un exemple un peu plus complexe, qui illustrera plusieurs éléments non présents dans l'exemple précédent. Le concept de ligne peut se définir comme suit (*la précision de la définition même n'est pas ici le point important*):

> *"Une ligne est un lieu géométrique composé de points qui sont tous contigus à un autre point ou à deux autres points. Deux points ne peuvent pas être contigus au même sous-ensemble de points."*



Note: *A partir de ce point, les éléments des ensembles ne sont plus représentés explicitement. Les noms des ensembles (I , A) sont à l'interieurs des rectangles représentant ces ensembles et le nombre d'éléments dans un ensemble est donné par "card(-nom de l'ensemble-)".*

Donc une ligne (*"I"*) a un nombre indéterminé de composantes qui sont des points. L'ensemble (*"A"*) de ces points est à la fois le domaine et le co-domaine de la projection "CONTIGU", donc ces points sont contigus entre eux. Il y a par contre des restrictions: un point n'est pas contigu à lui-même (*non-réflexivité*), il est contigu à un ou deux points (*ordre = <1..2>*) et deux points ne peuvent être contigus au même sous-ensemble de points (*fonction*); ils peuvent au plus partager un point contigu s'ils sont tous les deux contigus à deux points. Cette projection, enfin, est vérifiée pour l'intervalle de temps durant lequel l'identité de l'être est confirmée comme étant LIGNE (*temps = t0*).

On voit donc que les éléments du formalisme de description sont des concepts d'ensemble, de relation entre éléments d'ensembles, d'intervalles de temps durant lesquels ces relations sont vérifiées, et de descripteurs des propriétés de ces relations (*constituant en fait des méta-descripteurs: des descripteurs décrivant des descripteurs*).

Représentée de façon graphique (*voir ci-haut*), la définition d'un concept (*ou relation*) est un graphe où les nodes sont les ensembles et les arcs sont les projections. De plus, et cela deviendra apparent plus loin, une définition est représentée par un graphe connecté; c'est-a-dire qu'aucune partie du graphe n'est isolée des autres parties (*ceci constituerait DEUX graphes*).

* * * * *

A. Vers le formalisme

(DESCRIPTIONS PLUS OU MOINS PRECISES DE PAIRES ORDONNEES ETIQUETTEES ET TEMPORAIREMENT VALIDEES.)

Voici l'essence du formalisme de représentation. Nous allons maintenant tenter de transmettre la signification et les implications de cette formule synthétique au lecteur.

Quand un relation s'établit entre deux êtres, ceux-ci forment en quelque sorte une paire dont l'étiquette est le nom de la relation. Par exemple si Pierre est le père de Louise, nous avons la paire (*Pierre , Louise*) dont l'étiquette est PERE, ou si on veut : PERE (*Pierre Louise*) dans la notation propre à la logique propositionnelle.

La paire est ordonnée puisque "PERE (*Louise Pierre*)" n'est pas l'équivalent de PERE (*Pierre Louise*): l'ordre des éléments de la paire est donc important.

Lorsque nous faisons l'énoncé d'une relation entre deux êtres, nous limitons la validité de cet énoncé à une période de temps. Nous ne pouvons affirmer ORBITE-AUTOUR (*Terre Soleil*) sans délimiter un intervalle temporel de validité: "depuis près de 10 milliards d'années". Avant cette époque le système solaire n'était pas formé!

Toute affirmation ou énoncé est l'établissement de relations valides entre des êtres durant un intervalle de temps quelconque. L'aspect temporel est INDISSOCIABLE d'une relation.

Une affirmation est donc la description de paires ordonnées étiquettées et temporairement validées. Une affirmation peut être plus ou moins précise et comporter plus ou moins d'éléments d'information.

Par exemple, je peux affirmer en ce moment que "des planètes orbitent autour du Soleil" ou que "neufs planètes orbitent autour du Soleil" ou encore que "Mercure, Vénus, Mars, Jupiter, Saturne, Uranus, Neptune et Pluton orbitent autour du Soleil". Chaque affirmation décrit un ensemble de paires ordonnées. Dans le dernier cas chaque paire est individuellement formée :

ORBITE-AUTOUR (*Mercure Soleil*) / maintenant /
ORBITE-AUTOUR (*Venus Soleil*) / maintenant /
.
.
.
ORBITE-AUTOUR (*Pluton Soleil*) / maintenant /

Dans le deuxième cas, nous connaissons le nombre de paires et la classe à laquelle appartiennent tous les premiers membres des paires:

ORBITE-AUTOUR (*-une planète-Soleil*pn)/*maintenant*/
ORBITE-AUTOUR (-une 2ième planète-Soleil)/maintenant/
.
.
.
ORBITE-AUTOUR (*-une 9ième planète-Soleil*)/maintenant/

Enfin dans le premier cas, nous ne connaissons pas le nombre de paires, mais nous savons la classe à laquelle appartiennent tous les premiers membres des paires:

ORBITE-AUTOUR (*-une planète-Soleil*)/maintenant/
ORBITE-AUTOUR (*-une 2ième planète-Soleil*)/maintenant/
.
.
.
ORBITE-AUTOUR (*-une Nième planète-Soleil*)/maintenant/

Dans chacun des trois cas, nous connaissons l'identité exacte du deuxième membre de la paire. Or nous aurions pu faire une affirmation encore moins précise: "des planètes orbitent autour d'une étoile simple ou d'étoiles multiples". Maintenant aucune étoile spécifique (*comme le Soleil*) n'est nommée et même le nombre de soleils autour desquels orbitent des planètes n'est pas spécifié.

* * * * *

B. Concepts

En fait, cette description de paires ordonnées est suffisamment générale (*le nombre de paires n'est pas précisé, ni les individus qui y participent, seulement leurs types ou classes*) pour correspondre à n'importe quel système solaire. Cette description s'applique donc à plusieurs êtres mais pas à d'autres. L'ensemble des êtres auxquels elle s'applique forme une classe et cette classe nous lui donnons le nom (*ou étiquette*) de SYSTEME SOLAIRE.

Cet énoncé a donc une propriété extrèmement importante: il s'applique à un ensemble d'êtres et ne s'applique pas à un autre ensemble disjoint d'êtres. C'est presque la définition d'un concept. Il manque néanmoins quelque chose : l'information permettant d'identifier la surclasse du concept, c'est-à-dire la classe qui contient tous les exemples de notre concept et d'autres êtres en plus. Ajoutons-lui cette information:

> *"Un système (surclasse) dont les composantes sont une ou des planètes ainsi qu'une ou plusieurs étoiles et autour (de laquelle / de toutes lesquelles) orbitent toutes les planètes".*

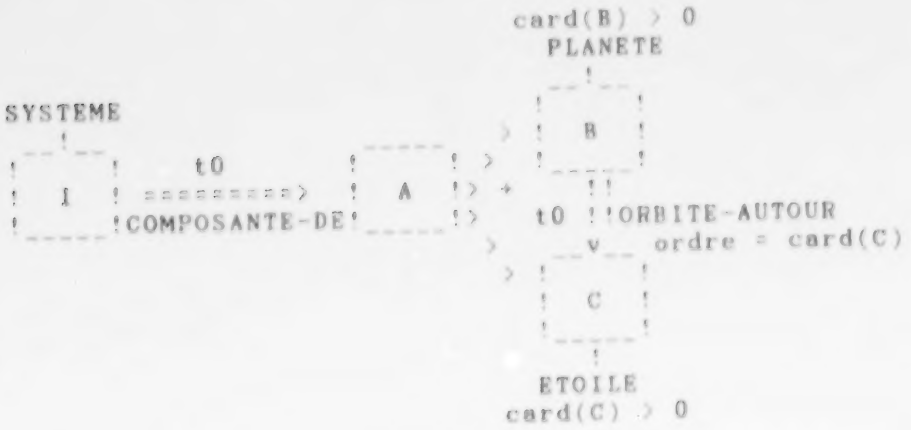
Nous avons ajouté le fait qu'un système solaire est une sorte de système avec des composantes particulières (*planètes et étoiles*) impliquées dans un processus particulier (*orbiter*). Un système solaire est donc une forme spécialisée de système.

Pour qu'une définition d'un concept soit complète, il faut d'abord donner la surclasse (*"genus" selon Aristote*) et ce qui fait du concept un classe restreinte de la surclasse (*"differentiae" toujours selon Aristote*).

Le "differentiae" constitue l'énoncé qui décrit des paires ordonnées, énoncé qui peut se représenter sous forme d'un graphe connecté.

Voici la représentation graphique de la définition du concept de SYSTEME SOLAIRE:

SYSTEME SOLAIRE (I) ::



Cet exemple introduit des éléments supplémentaires de description ainsi que leurs notations.

Pour que l'élément du simpleton "I" soit un SYSTEME SOLAIRE, il faut d'abord qu'il soit un SYSTEME. Ensuite, il doit pouvoir être décrit par des paires ordonnées dont la représentation graphique soit une forme plus précise que celle décrivant le concept SYSTEME SOLAIRE.

Dans la définition d'un concept (*ou d'une relation*), l'intervalle temporel t0 représente l'intervalle de temps durant lequel un être est identifié comme membre de la classe décrite par le concept (*ou durant lequel une paire ordonnée est considérée valide*).

Donnons d'abord la signification du graphe en le parcourant de gauche à droite: Un système solaire a un ensemble (*"A"*) dont tous les éléments forment chacun une paire ordonnée COMPOSANTE-DE avec ce système, i.e. chacun est une composante du système. Cet ensemble de composantes est le résultat de l'UNION des ensembles DISJOINTS "B" et "C" (*par l'opérateur "+"*).

L'ensemble "B" possède au moins un élément et chaque élément appartient à la classe PLANETE. L'ensemble "C" lui-aussi possède au moins un élément et chaque élément appartient à la classe étoile.

L'ensemble des planètes et des étoiles forment un ensemble de paires ordonnées nommées ORBITE-AUTOUR. Nous savons d'abord que tous les éléments de "B" et tous les éléments de "C" participent à la formation de paires (*par définition de projection*). Donc chaque planète orbite autour d'au moins une étoile.

Evidemment ceci n'est pas suffisant; chaque planète doit tourner autour de toutes les étoiles de l'ensemble "C". Donc en indiquant l'ORDRE de la projection (*le nombre de paires auxquelles participent chaque élément du domaine de la projection*) comme étant égal au nombre d'éléments du CO-DOMAINE, nous signifions que chaque planète orbite autour de chaque soleil. Autrement dit le nombre de paires ordonnées nommées ORBITE-AUTOUR égale le nombre d'éléments de "A" multiplié par le nombre d'éléments de "B".

Ce graphe est suffisamment complexe pour aborder le concept de paires connectées. La projection COMPOSANTE-DE du simpleton "I" sur l'ensemble "A" forment des paires ordonnées dont le co-domaine est "A" et la projection ORBITE-AUTOUR forme des paires ordonnées dont le domaine est un sous-ensemble du co-domaine de COMPOSANTE-DE. Donc des éléments du co-domaine d'une projection deviennent des éléments du domaine d'une autre projection.

EX:

ETIQUETTE1 (*el-1 el-2*) ETIQUETTE2 (*el-2 el-3*)

Voici deux paires ordonnées. Ces paires sont connectées par l'élément el-2. Nous pouvons COMPOSSER ces paires pour obtenir ETIQUETTE1*ETIQUETTE2 (*el-1 el-3*) par transitivité, ce qui signifie que el-1 est indirectement relié à el-3.

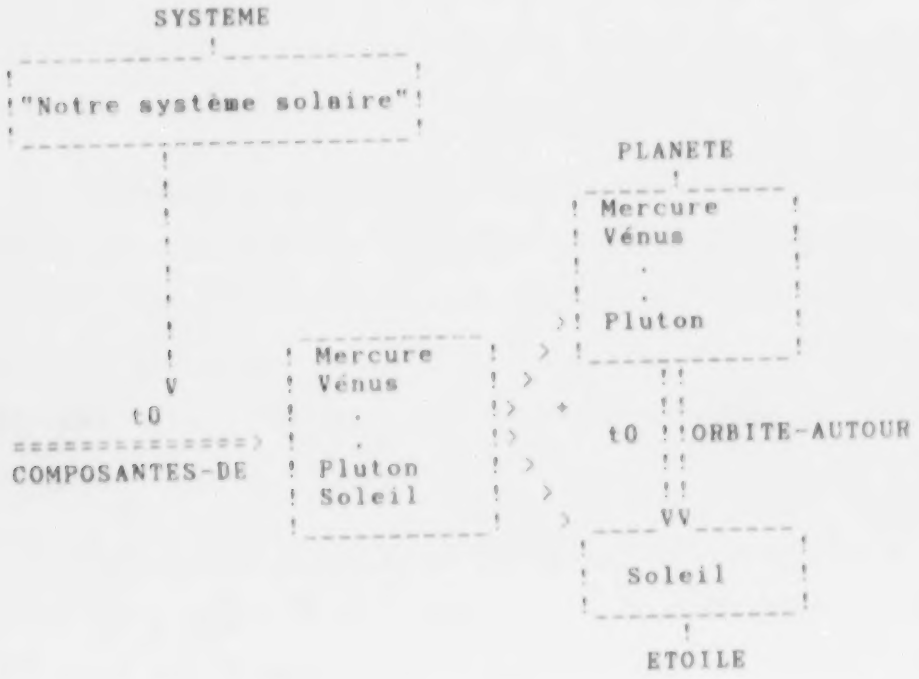
Dans un graphe définissant un concept, l'élément unique du simpleton "I" représente un être qui doit répondre aux critères énoncés par le graphe. Pour satisfaire à ces exigences, il doit former un ensemble de paires ordonnées telles que décrites avec plus ou moins de précision par le graphe. TOUS les éléments de ces paires seront alors directement ou indirectement reliés (*par composition*) à l'élément de "I", c'est-à-dire l'être dont l'appartenance à la classe est verifiée.

Ceci est intuitivement nécessaire; tous les éléments servant à décrire l'être en question doivent lui être directement ou indirectement reliés sinon ils sont sans pertinence. Ceci explique qu'un graphe définissant un concept se doit d'être un graphe connecté, c'est-à-dire un graphe où tous les éléments sont reliés à l'élément de "I".

Voyons maintenant de quelle façon la description de NOTRE système solaire satisfait aux exigences de la définition d'UN système solaire.

Décrivons NOTRE système solaire:

"C'est un système dont les composantes sont l'étoile Soleil et les planètes Mercure, Vénus ... et Pluton; Mercure, Vénus ... et Pluton orbitent autour du Soleil."



Note: *l'ordre de la projection ORBITE-AUTOUR est nécessairement égal à 1 qui est égal à card(<Soleil>). L'intervalle temporel t0 est celui durant lequel cette affirmation est faite.*

Nous devons vérifier que le graphe décrivant l'être "Notre système solaire" est conforme aux specifications du graphe définissant le concept SYSTEME SOLAIRE. Pour ce faire nous devons vérifier que le premier définit un ensemble de paires ordonnées répondant aux spécifications plus générales du deuxième.

Le graphe de "Notre système solaire" introduit des INDIVIDUS (*des êtres uniques et identifiés*) dans des ensembles délimités; tous les éléments d'un ensemble ainsi délimité doivent appartenir à la classe qui délimite l'ensemble.

La première action qui s'impose consiste à vérifier que "Notre système solaire" est bien un système. Pour cela nous devons vérifier que le graphe décrivant "Notre système solaire" est un cas particulier du graphe décrivant la classe SYSTEME (*nous verrons plus loin les mécanismes de vérification*).

Pour ensuite vérifier que Mercure, Vénus etc. sont biens des exemples de planètes, il faut accèder au graphe définissant PLANETE (*qui doit appartenir au réseau conceptuel*) et tester la description de chaque individu (*qui est un graphe*) sur le graphe de PLANETE. Finalement nous vérifions de façon similaire que Soleil est bien une ETOILE (*un autre concept dans le réseau*).

Nous venons donc de vérifier que tous les individus sont correctement des éléments des ensembles délimités. Il faut maintenant vérifier que les paires formées entre ces ensembles sont correctes. C'est-à-dire qu'entre autre les paires d'individus étiquettées ORBITE-AUTOUR répondent aux conditions validant cette relation.

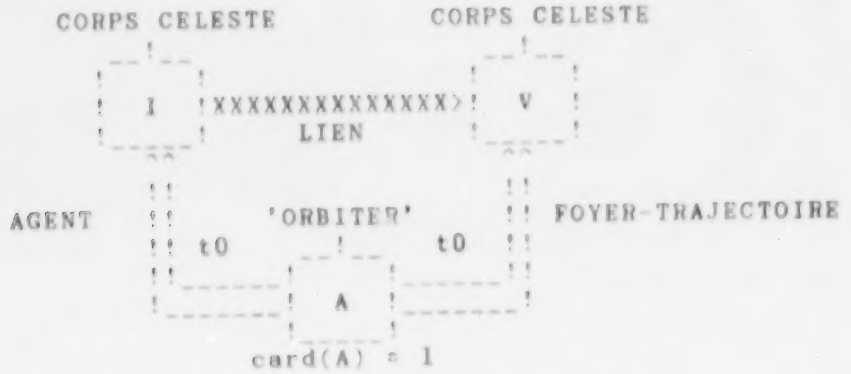
* * * * *

C. Relations

Qu'est-ce qui permet de valider ORBITE-AUTOUR (*Jupiter Soleil*)? La relation ORBITE-AUTOUR est, tout comme un concept, définie par un graphe donnant les conditions minimale nécessaires à l'établissement de la relation ORBITE-AUTOUR entre deux êtres. Cette relation est inscrite à l'intérieur d'un réseau relationnel tout comme chaque concept est inscrit dans un réseau conceptuel.

Examinons le graphe définissant la relation ORBITE-AUTOUR:

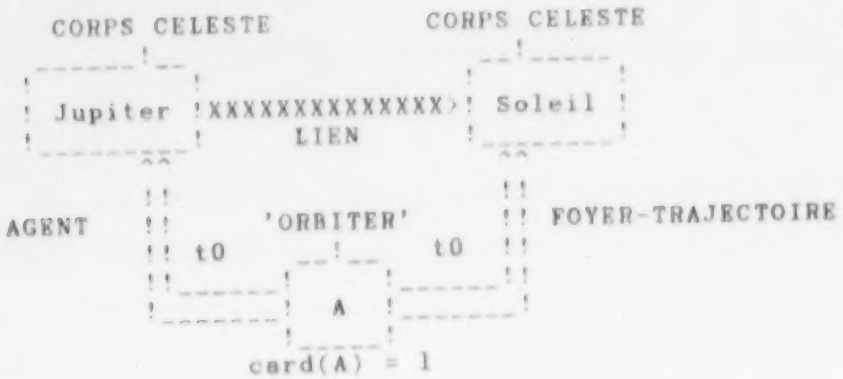
ORBITE-AUTOUR (I V) ::



Littéralement, ce graphe affirme que pour que la relation ORBITE-AUTOUR soit établie entre deux êtres, ces êtres doivent être chacun un CORPS CELESTE et doivent avoir la relation plus générale LIEN (*le "genus" ou surclasse relationnelle de ORBITE-AUTOUR*) entre eux.

Ensuite les deux êtres doivent répondre aux conditions plus spécifiques qui sont définies par le graphe: il y a une action de décrire une orbite (*succintement nommée par 'orbiter'*) dont l'AGENT est au moment de la validation de la paire (*t0*) le premier élément de la paire ordonnée et dont au même moment le foyer de trajectoire (*FOYER-TRAJECTORIEL*) est le deuxième élément.

Pour vérifier ORBITE-AUTOUR (*Jupiter Soleil*), il faut donc vérifier le graphe:



On doit démontrer qu'il y existe un cas de déplacement en orbite dont un corps céleste Jupiter est l'agent et un corps céleste Soleil est le foyer de la trajectoire. Il faut donc vérifier des appartenances à des classes et la validité de relations.

Donc pour vérifier l'appartenance de "Notre système solaire" à la classe SYSTEME, il a fallu vérifier l'appartenance d'individus à d'autres classes et vérifier des relations entre ces individus. De plus pour vérifier une relation entre deux individus, on doit vérifier l'appartenance d'individus à des classes et vérifier d'autres relations entre individus. On s'aperçoit donc que le processus est éminemment récursif.

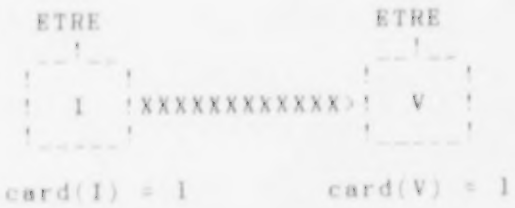
Pour que le processus s'arrête, il faut qu'éventuellement toute vérification mène à une vérification triviale (*immédiatement réussie*) d'appartenance à une classe ou de relation entre individus.

La surclasse commune de toutes les classes est décrite par le concept ETRE qui regroupe tous les individus. Cette classe n'a bien sur pas de surclasse. La surclasse commune de toutes les relations est la relation RELATION; elle est vraie automatiquement entre deux êtres. Il est toujours

possible de concevoir une relation entre n'importe quels deux êtres; cette possibilité est d'ailleurs documentée par le réseau relationnel lui-même. Cette relation n'a pas de surclasse relationelle:

ETRE (I) ::

RELATION (I V) ::



Pour que les réseaux conceptuel et relationnel permettent d'identifier des individidus et des relations entre individus, ils ne doivent en aucun cas comporter de cycles; un concept ou une relation ne peut être directement ou indirectement sa propre surclasse. De plus, bien sûr, une relation ne peut être la surclasse d'un concept et vice-versa.

Les réseaux permettent d'identifier des individus et des relations entre individus. La qualite des identifications dépend de la qualité des réseaux. Les réseaux peuvent être définis et modifiés. A tout moment les définitions de concepts et de relations doivent être cohérentes entre elles et les réseaux ne doivent pas contenir de cycles.

Nous allons maintenant aborder la description complète du formalisme de représentation de concepts et de relations ainsi que les contraintes de cohérence permettant de valider les réseaux eux-mêmes.

* * * * *

VI. DESCRIPTION COMPLETE DU FORMALISME DE REPRESENTATION.

A. PAIRES ORDONNEES, INTERVALLES TEMPORELS DE VALIDITE & GRAPHES.

Une paire ordonnée est composée de deux éléments ayant entre eux une relation définie, relation dont la validité n'est affirmée que pendant un intervalle temporel donné. Cet intervalle peut être completement spécifié (*le 15 août 1985 à 16:00 heures, heure de Greenwhich*) ou décrit relativement à un autre intervalle temporel (*avant, juste avant, pendant, etc.*).

L'élément d'une paire peut être un individu (*"le vase chinois #1432"*), une paire ordonnée d'éléments ou un graphe (*qui est un ensemble plus ou moins précisément determiné de paires ordonnées connectées*). La définition d'élément est donc récursive avec le cas de base étant l'individu.

Un graphe est l'interprétation formalisée d'une description: description d'une situation précise (*schéma*), d'une classe (*concept*) ou des conditions de validation de paires ordonnées (*relation*).

Une interprétation d'un graphe est un ensemble de paires ordonnées qui répondent aux spécifications représentées par le graphe. Un graphe peut avoir un ensemble plus ou moins vaste d'interprétations dépendamment de la généralité de ses spécifications. Le graphe définissant le concept ANIMAL représente plusieurs interprétations. Idéalement tous les schémas d'animaux (*description de la chèvre de mon voisin etc.*) seront des interprétations valides du graphe en question.

Si un graphe détermine un ensemble d'interprétations qui est inclus dans l'ensemble des interprétations possibles du graphe ANIMAL, alors ce premier graphe décrit une sous-classe du concept ANIMAL. Il est possible alors d'expliciter cette relation de sous-classe en indiquant dans le graphe sa surclasse et en ne préservant que les spécifications non contenues dans la surclasse. Ainsi peut s'organiser une hiérarchie conceptuelle.

La même discussion s'applique parallèlement aux relations.

* * * * *

B. PROJECTIONS & DESCRIPTEURS

Il est possible de décrire un ensemble de paires ordonnées en décrivant le domaine (*l'ensemble des premiers éléments des paires*), le co-domaine (*l'ensemble des deuxièmes éléments des paires*), les contraintes définies que les éléments respectent (*relationpn*) *et l'intervalle temporel de validité des paires.*

L'ensemble des descriptions d'un ensemble de paires ordonnées forme une PROJECTION. Alors que des descripteurs s'appliquent aux ensembles que sont le domaine et le co-domaine, d'autres décrivent certaines propriétés des paires elles-mêmes (exemple: combien de fois un élément du domaine participe à la formation de paires).

Un ensemble d'éléments (*domaine ou co-domaine d'une projection*) peut être restreint dans l'identité de ses éléments soit par inclusion obligatoire à une classe ou par exclusion obligatoire d'une classe.

Exemple d'INCLUSION:

signifie que chaque élément de l'ensemble A est un animal.

Exemple d'EXCLUSION:

signifie qu'aucun élément de l'ensemble A n'est pas un animal.

Il est possible de décrire le nombre d'éléments d'un ensemble.

signifie un ensemble de 4 à 9 animaux.

$3 < card(A) < 10$

signifie un ensemble infini de points

$card(B) = infini$

signifie un ensemble C composé d'un nombre quelconque de pattes.

$card(C) = indéterminée$

Certaines propriétés d'un ensemble de paires connectées formées d'un domaine et de co-domaine peuvent faire l'objet de description.

Si chaque élément du co-domaine ne forme de paire qu'avec un élément du domaine, la projection est une FONCTION. Si l'inverse est vrai, nous avons une FONCTION INVERSE.

Si chaque élément du domaine forme des paires avec un sous-ensemble particulier du co-domaine (*aucun autre élément du domaine ne projète sur le même sous-ensemble du co-domaine*), nous avons une projection qui est une QUASI-FONCTION. Si l'inverse est vrai, nous avons une QUASI-FONCTION INVERSE.

On peut spécifier l'ORDRE d'une projection : description du nombre de paires auxquelles les éléments du DOMAINE participent. L'ORDRE INVERSE donné le nombre de paires auxquelles les éléments du CO-DOMAINE participent.

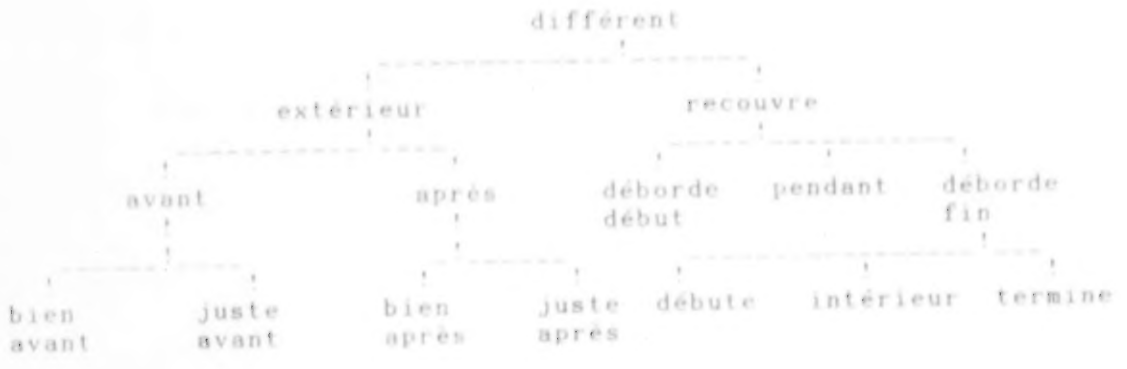
A chaque projection est attaché un intervalle temporel de validité. Cet intervalle peut être complètement spécifié ou peut être mis en rapport à un autre intervalle temporel.

Voici une liste des descriptions relatives applicables:

NOTE: *t1 et t0 représentent des intervalles temporels:*
>>>>>>>> illustre spatialement l'étendue de t1,
//////// l'etendue de t0,
XXXXXXX l'intersection de t1 et t0

		exemple
1.	t1 avant t0	>>>>>>>>> //////////
2.	t1 immédiatement avant t0	>>>>>>>>>//////////
3.	t1 après t0	//////// >>>>>>>>>>
4.	t1 immédiatement après t0	//////////>>>>>>>>>
5.	t1 pendant t0	////XXXXXXXXX//////
6.	t1 recouvre t0	>>>>>>>XXXXXXX///////
7.	t1 déborde le début de t0	>XXXXXXX/////////
8.	t1 déborde la fin de t0	///////////XXXXX>>>>
9.	t1 débute t0	XXXXXXXXXX//////
10.	t1 termine t0	///////////XXXXXXX
11.	t1 extérieur à t0	///////// >>>>>>>>>
12.	t1 intérieur à t0	>>>>>>>XXXXXXXXX>>>>
13.	t1 différent de t0	////XXXXXXXX>>>>>>>>
14.	t1 bien avant t0	>>>>>>>>> ////////
15.	t1 bien après t0	////////// >>>>>

Organisation en arborescence des descriptions relatives d'intervalles temporels:



Les ensembles peuvent être décrits par partition de leurs éléments, partition qui sépare les éléments d'un ensemble en deux sous-ensembles. Deux types de partitions permettent deux degrés de précision dans un énoncé.

La PARTITION STRICTE sépare un ensemble d'éléments en deux sous-ensembles disjoints.

Signifie : *"Parmi cinq étendues d'eau, certaines sont salées et d'autres non-saléesp".*

La PARTITION FLOUE sépare un ensemble en sous-ensembles qui ne sont pas nécessairement disjoints.

Signifie: *"Parmi plusieurs personnes, plusieurs se connaissent"*

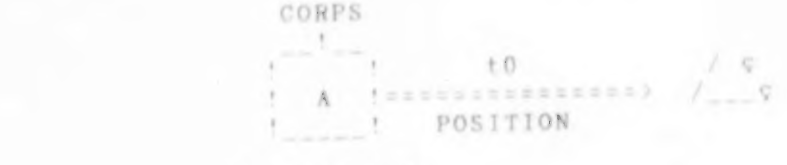
Note: *Une personne se connaît elle-même!*

* * * * *

C. RELATIONS SPECIALES

Dans un univers continuellement en changement, bien des phénomènes sont trop instables pour pouvoir être décrit de façon discrete (*au sens mathématique du terme*). Par exemple, pendant un intervalle temporel un corps peut occuper succesivement plusieurs positions; on ne peut donc former une paire ordonnée décrivant sa position précise lors de cet intervalle.

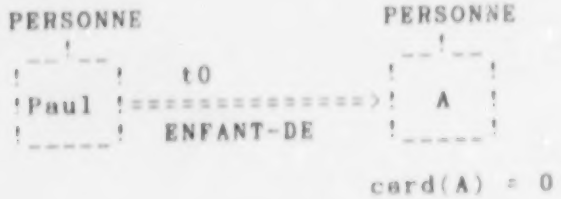
Pour représenter l'aspect instable d'un attribut d'un être, un type spécial de relation doit être défini. Une RELATION FLOUE indique que durant un intervalle temporel, plus d'une valeur peut être reliée à un être.



card(A) = 1

Signifie: *"Un corps occupe plusieurs positions lors de son observation."*

Une autre relation spéciale indique l'impossibilité de former des paires ordonnées. Si la cardinalité du co-domaine est zéro, alors aucune paire ne peut être formée.



Signifie: *"Paul n'a pas d'enfant."*

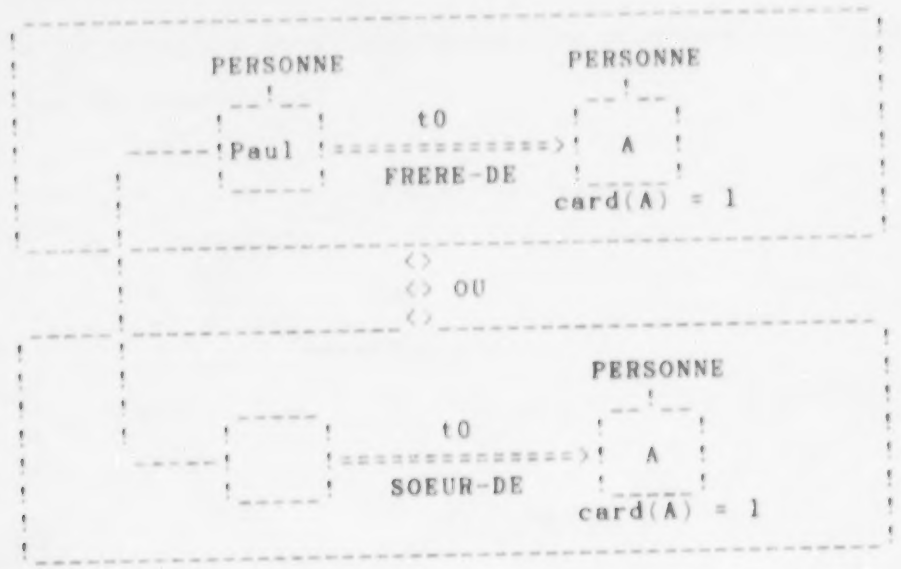
* * * * *

D. CONNECTEURS LOGIQUES, REGIONS D'UN GRAPHE & LIGNES D'IDENTITE

Une affirmation est représentée par un graphe qui spécifie un ensemble de paires ordonnées connectées. Parfois une affirmation peut proposer soit une disjonction d'interprétations soit une implication entre deux interprétations. Le graphe comporte alors deux régions, chacune spécifiant un ensemble de paires ordonnées connectées. Les deux régions peuvent alors n'être connectées que par les connecteurs logiques. Les régions étant distinctes, si un ensemble est partage par les deux régions alors cet ensemble n'est complètement décrit que dans une région. Sa présence dans l'autre région est indiquée par une ligne d'identité. Les exemples suivants clarifieront ces énoncés:

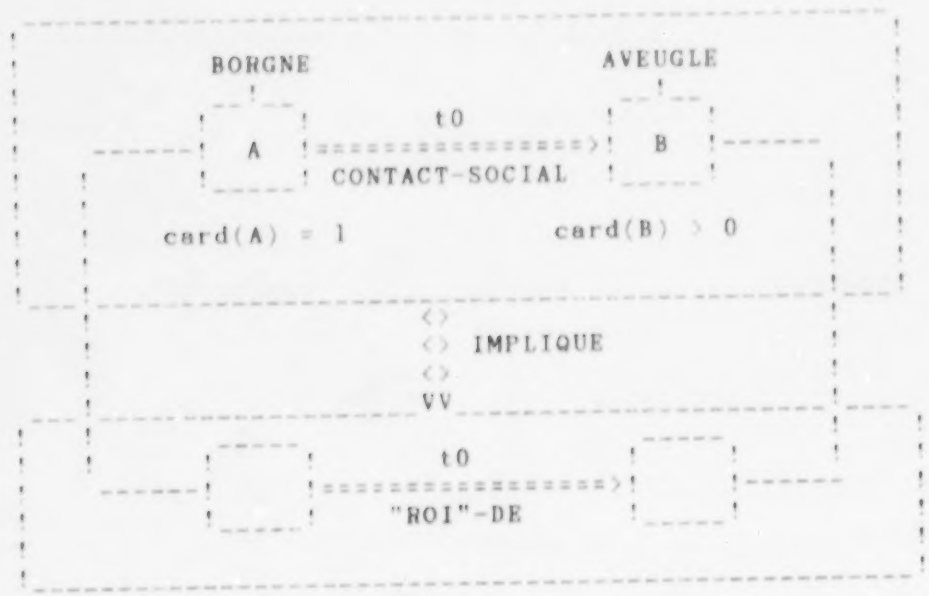
"Paul a soit un frère ou soit une soeur."

Cette affirmation contient une disjonction (*OU*). Voici une représentation graphique de cette affirmation:



"Si un borgne est parmi des aveugles alors il est leur roi."

Cette affirmation contient une implication. Autrement dit, si il y a un ensemble de paires ordonnées répondant aux spécification de l'antécédent de l'affirmation alors il doit y avoir un ensemble de paires ordonnées répondant aux spécifications du conséquent:

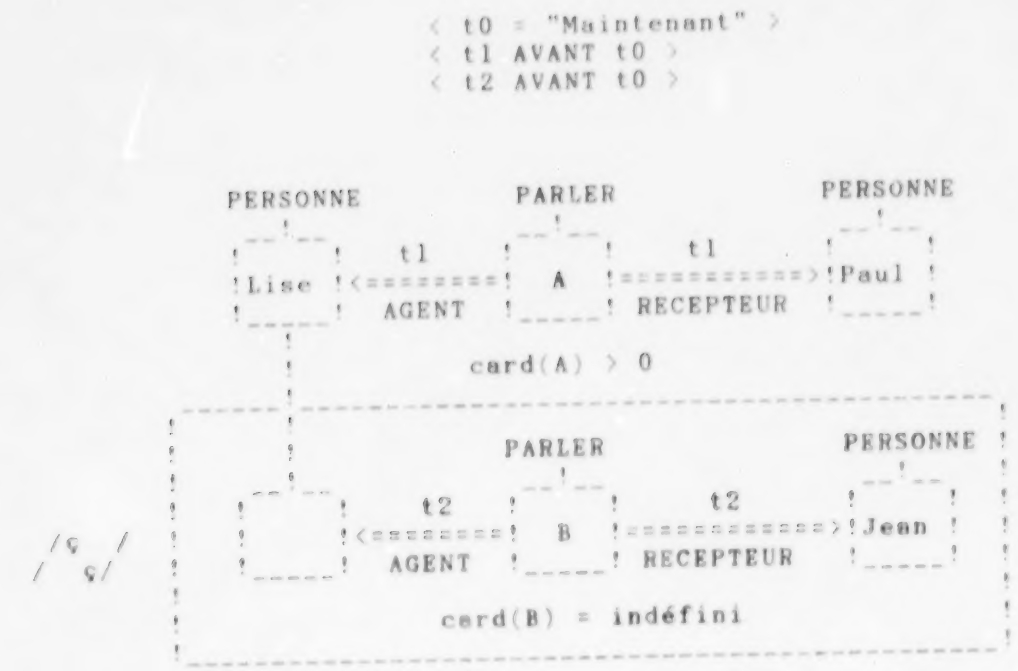


Nous avons vu auparavant comment indiquer qu'un ensemble non-connecté de paires ordonnées (*une projection*) était impossible, ce qui est une forme de négation.

Une autre forme de négation s'adresse à un énoncé dont la représentation est un ensemble de paires ordonnées connectées (*prescrites par un graphe*). La négation s'appliqué au graphe en entier ou à une région du graphe. Cette négation signifie qu'aucune interprétation valide n'est possible pour la région du graphe qui est niée:

"Lise n'a jamais parlé à Jean mais elle a parlé à Paul".

Voici la représentation de cet énoncé:



Notons de quelle façon les intervalles temporels sont décrits: t0 fait référence au moment ou l'énoncé est émis, t1 et t2 sont situés avant t0 sans plus de précision.

* * * * *

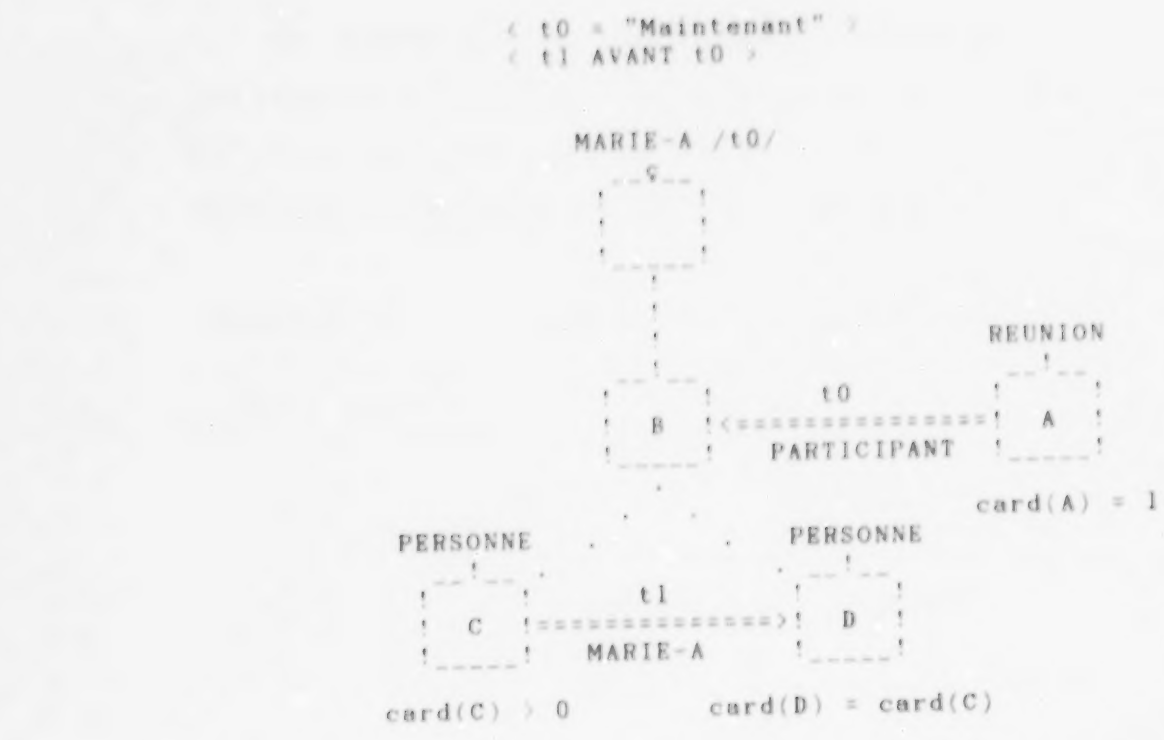
E. PAIRES ORDONNEES & GRAPHES COMME ELEMENTS DE PAIRES ORDONNEES

Un ensemble de paires ordonnées est spécifié par un domaine, un co-domaine, un intervalle temporel et une relation imposant des contraintes définies sur la formation de paires, tout ceci constituant une projection.

Les éléments du domaine et du co-domaine peuvent être autre chose que des individus; ils peuvent être des paires ordonnées (*une paire ordonnée formée de paires ordonnées est concevable*) ou des graphes.

Voyons d'abord un exemple d'un domaine d'une projection dont les éléments sont des paires ordonnées:

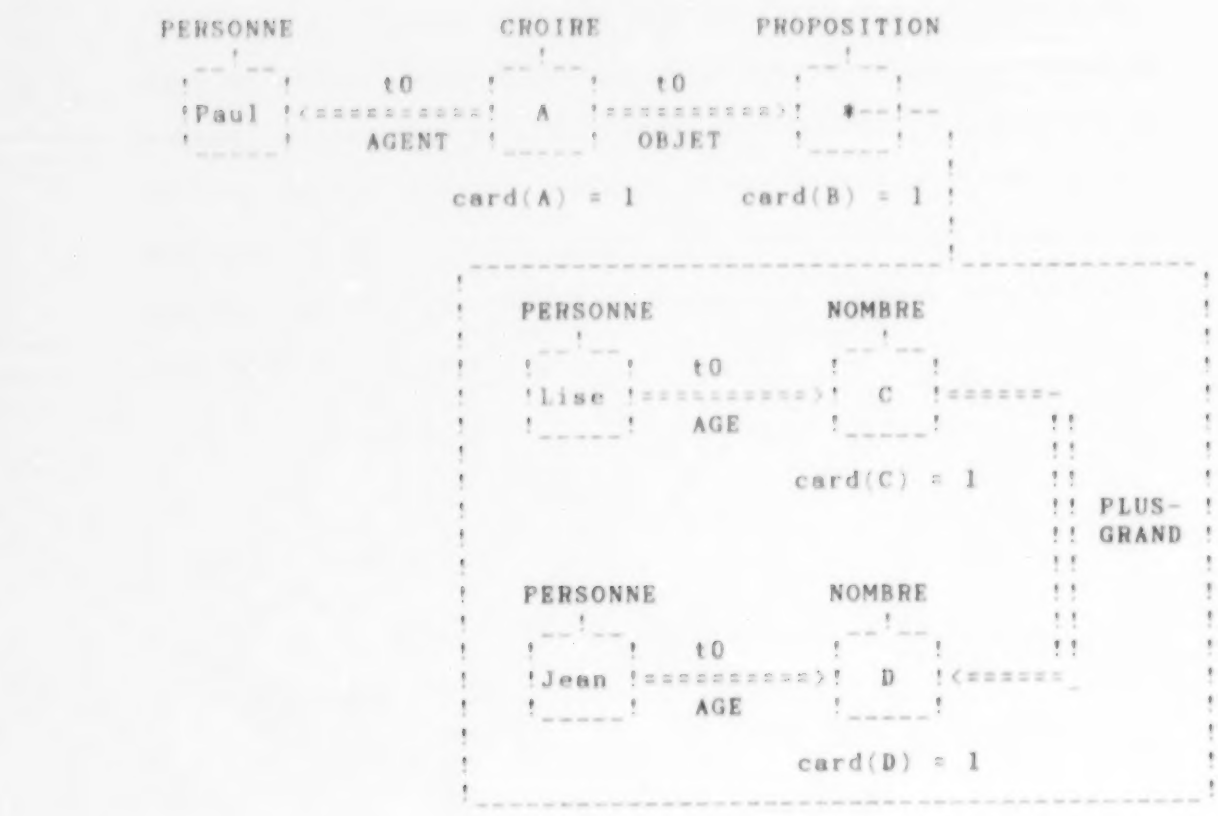
"Une réunion de couples divorcés"



Les lignes pointillées menant à l'ensemble B indiquent la provenance des paires ordonnées qui sont les éléments de cet ensemble, en l'occurence de la projection du domaine C au co-domaine D par la relation MARIE-A durant un intervalle temporel t1 précédent celui dans lequel l'affirmation est faite (*t0*).

Notez aussi l'emploi d'une ligne d'identité pour soumettre une copie de l'ensemble de paires ordonnées B à un test d'exclusion; celui-ci prescrit qu'aucune paire ne valide la relation MARIE-A au moment t0 de l'affirmation.

Un graphe peut être un élément d'un ensemble décrivant une projection. Prenez l'affirmation suivante: "Paul croit que Lise est plus agée que Jean". Ceci ne signifie pas qu'il connaît l'age exact de Lise ou de Jean:



Notez que la ligne d'identité relie le graphe à l'astérisque à l'intérieur de la boite symbolisant un ensemble: ceci signifie que le graphe est un élément de cet ensemble.

Ce que Paul croit est une proposition qui spécifie un ensemble de paires ordonnées connectées. Un de ces ensembles pourrait être :

< AGE(Lise 20)/t0/ AGE(Jean 19)/t0/ >

Dans ce cas-ci, la réalité ne contredit pas la conviction de Paul. Paul pourrait neanmoins se tromper. Paul ne connaît pas un ensemble de paires ordonnées, il croit que de telles paires doivent répondre à des spécifications exprimées par un graphe. Un graphe n'est pas la MEME chose qu'un ensemble de paires ordonnées connectées; c'est une spécification d'un tel ensemble.

* * * * *

F. OPERATIONS DE SPECIALISATION & DE GENERALISATION

Nous avons maintenant établi dans une mesure acceptable la sémantique de la notation permettant de décrire des situations en termes relationels et conceptuels, de définir des concepts et des relations.

Comme il en a été question précédemment les concepts et les relations forment des réseaux par inter-référentialité (*dimension horizontale*) et par subordination (*dimension verticale*). Un réseau doit être cohérent sur ses deux dimensions. Nous verrons plus loin comment se vérifie la cohérence co-référentielle. Voyons d'abord comment se vérifie la cohérence de l'organisation hiérarchique des concepts et des relations.

Lors de la définition d'un concept ou d'une relation, on doit donner la surclasse de ce concept ou de cette relation afin de l'inscrire dans la dimension verticale du réseau. La question est de vérifier si cette surclasse est correcte.

NOTE: *La discussion suivante traitera seulement des concepts; le même traitement peut être appliqué aux relations.*

Un concept C1 est une sous-classe d'un concept C0 si l'univers des interprétations valides du graphe de C1 est inclus dans l'univers des interprétations valides du graphe de C0. Comment peut-on s'assurer de cette relation?

Prenons une projection P. Introduisons une opération S sur une projection dite de spécialisation ayant la propriété suivante:

S(P) ---> P (implication logique)

(où S(P) (l'application de S sur P) produit une nouvelle projection P' telle que P' implique P. Autrement dit P' est un cas spécial de P).

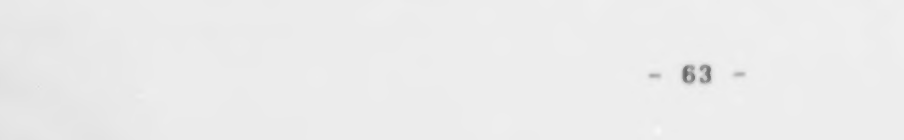
L'opération de généralisation G est l'opération inverse de la spécialisation:

P ---> G(P) (*implication logique*)

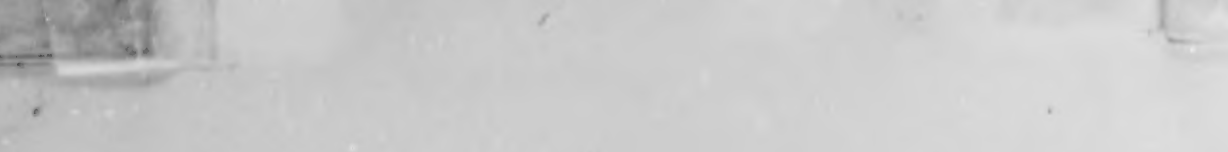
Un concept C1 est une sous-classe d'un concept C0 si le graphe de C1 est équivalent à celui de C0 dont des projections auraient été spécialisées et/ou auxquel des projections auraient été ajoutées.

Le graphe de C1 peut être représenté de la façon suivante: on indique que C0 est la surclasse de C1 et on inscrit les projections spécialisées dans C0 et les projections ajoutées à C0.

TRIANGLE (I) ::



TRIANGLE SCALENE (I) ::



"Un triangle dont les côtés ont chacun une longueur différente"

Le simpleton I indique (*par opération d'inclusion*) que tout TRIANGLE SCALENE doit d'abord être un TRIANGLE.

Ici, TRIANGLE SCALENE est le résultat de l'ajout d'une projection dont la relation est LONGUEUR. Pour connecter cette projection au simpleton I, il a fallu re-inscrire la projection de I sur A. Il n'a pas été nécessaire de re-inscrire le descripteur de cardinalité de l'ensemble A.

Il faut donc lorsqu'on inscrit une projection spécialisée ou ajoutée la connecter directement ou indirectement au simpleton I; ceci nécessitera parfois de "recopier" des projections non-modifiées provenant de la surclasse. Pour de telles projections le minimum d'information nécessaire à leur identification sera donnée (*domaine, relation, intervalle temporel et co-domaine*).

Ce minimum d'information est suffisant pour identifier une projection dans la lignée au-dessus d'un concept à cause de la REGLE D'EXCLUSION MUTUELLE des projections ayant le même domaine, règle que nous introduisons ici:

SI P1 et P2 sont des projections d'un graphe ayant le même domaine et décrits respectivement par des intervalles temporels t1 et t2 tels que t1 n'est pas EXTERIEUR à t2

ET SI r1 et r2 sont respectivement les relations de P1 et P2

ALORS r1 est différent de r2
et r1 n'est pas une sous-classe relationnelle de r2
et r2 n'est pas une sous-classe relationnelle de r1.

La surclasse d'où provient la projection qui est spécialisée n'est pas nécessairement la surclasse DIRECTE du concept; il peut s'agir d'une surclasse indirecte. Dans ce cas la règle suivante doit être respectée:

SI C0 est une surclasse indirecte de C1

ALORS C1 peut contenir une projection spécialisée de C0 SI ET SEULEMENT SI cette projection n'a pas été déjà spécialisée dans la lignée entre C0 et C1.

La question qui se pose est de savoir comment on spécialise une projection.

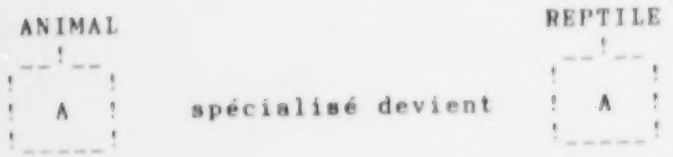
On spécialise une projection en AUGMENTANT LA PRECISION des descripteurs.

On peut donner une description plus précise des éléments d'un ensemble en les NOMMANT (*si ce sont des individus*) ou en les spécialisant (*si ce sont des paires ordonnées ou des graphes*).

On spécialise une paire ordonnée en utilisant une sous-classe de la relation l'identifiant ou en précisant l'intervalle de validité.

On spécialise en précisant la cardinalité d'un ensemble (*card(A) > 3 devenant card(A) > 5 ou encore card(A) = 10*).

On peut aussi rendre le test d'inclusion plus sévère en utilisant une sous-classe du concept utilisé:



On peut rendre le test d'exclusion plus sévère en utilisant une surclasse du concept utilisé (*i.e. par généralisation d'une négation*).

On peut spécialiser une projection en remplaçant la relation par une sous-classe de cette relation (*la validité de l'opération dépend de la nature du réseau relationnel*).

Finalement en précisant l'intervalle temporel et les propriétés de la projection (*ordre, fonction etc.*) on obtient une projection spécialisée.

Lorsque le graphe contient des connecteurs logiques reliant des régions du graphe, il est possible de spécialiser le graphe en spécialisant ou en généralisant une ou des régions selon les cas.

Prenons chaque connecteur logique l'un après l'autre. Region1 et Region2 seront les régions du graphe connectées. Voyons comment on peut spécialiser le graphe:

Region1 OU Region2 : -*spécialiser Region1 ou Region2.*

-*faire disparaitre une région (Region1 ---> (Region1 OU région2))*

NEGATION (Region1) : -*généraliser une projection dans Region1*

Region1 IMPLIQUE Region2 : *ceci est logiquement equivalent a NEGATION (Region1) OU Region2 donc*

- *généraliser Region1 ou spécialiser Region2*
- *faire disparaitre Region2 et garder NEGATION (Region1)*
- *faire disparaitre Region1*

On spécialise une région en spécialisant au moins une de ses projections. On généralise une région en généralisant au moins une de ses projections.

Nous pouvons maintenant émettre la règle inductive suivante:

Un concept C1 (ou relation R1) est une sous-classe du concept C0 (ou relation R0) SI ET SEULEMENT SI

C1 (ou R1) est le résultat d'une opération de spécialisation sur C0 (ou R0)

OU

C2 (ou R2) est une sous-classe de C0 (ou R0) ET C1 (ou R1) est le résultat d'une opération de spécialisation sur C2 (ou R2).

On peut concevoir la DISTANCE entre un concept (*ou relation*) et une surclasse comme le nombre maximal d'opérations de spécialisation qui transformerait la surclasse en sa sous-classe.

Cette distance est dépendante de la nature des réseaux: Si de nouvelles sous-classes sont ajoutées par exemple entre la relation R12 et sa sousclasse R13, remplacer R12 par R13 dans une projection n'est plus au plus une seule spécialisation mais théoriquement plusieurs.

* * * * *

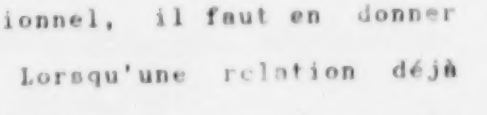
VII. DYNAMISME DE LA CREATION & DE LA MODIFICATION DES RESEAUX (*ANALYSE DE CONCEPTS ET DE RELATIONS*)

Essentiellement créer des réseaux consiste à spécialiser le concept et la relation fondamentales: ETRE et RELATION. La première spécialisation possible pour ETRE est bien sur d'ajouter une projection puisque ETRE n'en contient pas.

ETRE (I) ::

card(I) = 1

RELATION (I V) ::



Lorsqu'une projection est ajoutée contenant une relation non-présente du réseau relationnel, il faut en donner la définition dans ce réseau. Lorsqu'une relation déjà définie est employée, elle doit l'être en accord avec sa définition; les éléments du domaine et co-domaine de la projection l'utilisant ne doivent pas contredire les contraintes énoncées dans la définition de la relation.

Faire l'analyse de concepts est l'équivalent de construire des réseaux (*conceptuel et relationnel*) cohérents ou chaque graphe capture la sémantique d'un concept ou d'une relation telle que perçue par une communauté d'êtres pensants. C'est un acte créatif rigoureux et difficile.

Un système d'aide à l'analyse de concepts (*et relations*) s'avèrerait donc un outil très utile. Pour qu'un tel système puisse exister, il faut qu'il ait la capacité théorique de guider et d'orienter les efforts d'analyse de son utilisateur.

Imaginons abstraitement l'acte de définir un concept. Il existe un nombre infini de graphes qui pourraient représenter ce concept. Un nombre moins grand de graphe répondrait aux contraintes de cohérence produites par les réseaux. Encore moins de graphes, en plus, représenteraient de façon acceptable la sémantique du concept.

Plus les réseaux sont complexes et larges, plus les contraintes de cohérences qu'ils imposent à un graphe nouveau sont nombreuses; autrement dit la marge de décision est de plus en plus mince.

Ceci indique qu'il est préférable d'ensemencer les réseaux de relations et de concepts généraux et fondamentaux afin de creer dès le début une situation où les choix sont relativement limités lors de la définition d'un nouveau concept ou relation.

Le rôle d'un système d'aide à l'analyse de concepts consiste surtout à rendre explicite et compréhensible ces contraintes de cohérence afin de diminuer pour l'utilisateur la complexité des choix qu'il doit prendre. Par contre, en aucun cas, le système est-il capable de valider la justesse sur le plan sémantique d'une définition. Il ne peut qu'en vérifier la validité structurelle.

Le système pourra conseiller l'emplacement d'un concept (*ou relation*) dans le réseau en optimisant l'aspect hiérarchique : le réseau doit à chaque moment contenir le plus de liens de surclasse possible.

Le système devra de plus faire la liaison entre le formalisme de représentation (*permettant des manipulations automatiques*) et un dialogue en langage naturel (*avec ses ambiguités, ses aphorismes etc.*).

Il reste à découvrir les heuristiques diverses qui seront responsable des différents aspects du comportement du système d'aide à l'analyse de concepts. Elles s'exprimeront en termes des propriétés structurelles des réseaux représentés par le formalisme qui a été mis au point.

* * * * *

VIII. ARC 0.1, UN PRE-PROTOTYPE ET PROSPECTIVES

Afin d'illustrer les propriétés d'un réseau conceptuel et de démontrer la possibilité pour un programme de faire du raisonnement symbolique sur des définitions de concepts, ARC 0.1 a été conçu.

Les concepts sont définis en utilisant un sous-ensemble restreint du formalisme de représentation.

Le logiciel propose à son utilisateur d'insérer dans un réseau hiérarchique différents concepts de géométrie (*polygones*) pré-définis (*les définitions sont accessibles à l'utilisateur*). Un concept est inséré en lui donnant une surclasse déjà présente dans le réseau (*au début le concept POLYGONE est déjà inscrit*). ARC 0.1 accepte l'insertion, la refuse ou la précise.

Si l'insertion est refusée, ARC 0.1 explique pourquoi en termes des définitions respectives du concept et de sa surclasse proposée. Si l'utilisateur a donné une surclasse correcte mais indirecte (*au moment du choix*), ARC 0.1 "descend" le concept à sa position correcte.

L'exercice est terminé lorsque tous les concepts prédéfinis ont reçu toutes leurs surclasses possibles.

Relativement peu de temps a été alloué à la mise au point de ce pré-prototype. L'effort majeur fut porté sur le développement du formalisme.

ARC 0.1 est écrit en LISP (*dialecte IQLISP version 1.7*) pour le IBM PC et requiert au moins 256K de mémoire pour fonctionner. Le programme est interprété et occupe sous forme non-étendue (*avant l'expansion des fonctions macros*) près de 50K de mémoire.

Il présente le réseau en construction de façon graphique et les options disponibles sont organisées en menus interactifs à l'épreuve des erreurs d'entrée.

La réalisation d'un système d'aide complet à l'analyse des concepts demandera des ressources informatiques de développement considérables; au moins une station de travail pour l'Intelligence Artificielle: Tektronik serie 4400, Xerox série 1100 ou SUN II avec environnement POPLOG sont proposés. La préférence va à l'environnement POPLOG qui permet aux programmes d'être exprimés en partie en PROLOG (*langage déclaratif*) et en partie en POP-11 (*langage fonctionnel plus puissant que LISP*).

Le système cible (*présent dans les écoles*) serait de préférence du type "Mac-intosh" pour ses qualités d'interface et sa puissance computationnelle. Le IBM PC et les compatibles seraient un deuxième choix. L'utilisation d'un langage déclaratif tel PROLOG sera presqu'une nécessité étant donné le type de solutions associées au traitement de réseaux (*nécessité d'un backtracking automatique, programmation prescriptive etc.*). Pour rendre les programmes en Prolog performants, l'emploi d'un Prolog compilé sera nécessaire (*ARITY PROLOG pour le IBM PC et MacProlog pour le Macintosh*).

Le développement d'un formalisme de représentation des concepts est l'étape première et fondamentale de tout système informatique espérant traiter de l'information conceptuelle de façon adaptive et intelligente. Cette étape est en voie d'être franchie.

La mise au point d'un système d'aide à l'analyse des concepts est maintenant une entreprise théoriquement possible et déjà bien amorcée. Pour la mener à bien dans des conditions acceptables et réalistes, il faudra néanmoins compter sur un effort soutenu d'une durée que l'on peut estimer à deux ans.

* * * * *